

SPECIFICATION

WEB PAGE UPDATE NOTIFICATION METHOD AND WEB PAGE UPDATE NOTIFICATION DEVICE

FIELD OF THE INVENTION

本発明は、ネットワーク上で公開されるWebページが更新されたことを通知する更新通知方法及び装置に関する。

BACKGROUND OF THE INVENTION

近年、インターネットの普及とともに、多くのWebサイトが公開され、企業から個人に至るまで、これらWebサイトにより多彩な情報を公開している。

これらWebサイトのWebページの更新は、公開している各運営者の都合により行われるため不定期なものとなる。このため利用者は、Webページに頻繁にアクセスしなければ、更新を直ちに知ることはできない。そして、更新されたか否かを確認するためにアクセスしたにもかかわらずWebページの情報は古いままでいることも多く、無駄なアクセスを強いられる場合もある。

従来技術である特開2002-73455号公報には、仲介サーバによりWebページの事前チェックを行い、更新があったWebページのタイトルとURL(：Uniform Resource Locator)とを利用者に電子メールで通知する装置が開示されている。利用者に通知される電子メールの例を図4のM1に示す。か

かる装置によれば、利用者は、更新があった旨電子メールで通知されたW e bサイトのW e bページにアクセスすることで、無駄なアクセスを防止できる。

しかし、上記従来技術では、通知される情報が、更新のあったW e bページを特定する情報（U R L等の特定情報）にとどまっており、更新の内容を把握する手段がなかった。そのため利用者は、W e bサイト及びW e bページに実際にアクセスし、利用者自身にとって有意な情報の更新であるか否かを確認する必要があった。本発明は、以上の問題点に鑑みてなされたものであり、その目的は、利用者にとって有意な内容においてW e bページの更新があったか否かを容易に判断し得る更新通知を提供可能な更新通知装置及び方法を提供することにある。

SUMMARY OF THE INVENTION

上記課題を解決し、利用者にとって有意な内容のW e bページの更新があったか否かを容易に判別し得る更新通知を出力する更新通知装置及び方法を実現するために、本発明は、予め設定されたアドレスにより識別される少なくとも1つのW e bサイトに繰り返しアクセスし、該W e bサイトにおいて更新されたW e bページがある場合に、その更新の有無を含む通知文を外部に出力する更新通知装置及び方法であり、該更新されたW e bページから、更新内容を抽出し、該通知文に該更新内容の見出しと該更新内容との何れか若しくは両方と、該W e bページのアドレスとを付加して出力する。

BRIEF DESCRIPTION OF THE DRAWINGS

図 1 は、本発明の第 1 の実施例を示したブロック図であり、更新通知装置を含むネットワーク構成全体が示されている。

図 2 は、図 1 に示された更新通知装置における更新通知の処理手順を示したフローチャートである。

図 3 は、図 2 に示された処理手順において、差分情報を取得する方法を示した説明図である。

図 4 は、更新通知装置から利用者端末に送信される通知文の例として、従来の通知メール M 1 と本発明を適用したメールマガジン M 2 を対比して例示した図である。

図 5 は、本発明の第 2 の実施例を示したブロック図であり、更新通知装置を含むネットワーク構成全体が示されている。

図 6 は、図 5 に示された更新通知装置における更新通知の処理手順を示したフローチャートである。

図 7 は、図 5 に示された構成において用いられるテンプレート及びメールの例を示した図である。

図 8 は、本発明の第 3 の実施例を示したブロック図であり、更新通知装置を含むネットワーク構成全体が示されている。

図 9 は、図 8 に示された更新通知装置における更新通知の処理手順を示したフローチャートである。

図 10 は、図 8 に示される構成において出力される通知文としてのメールの例を示した図である。

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

本発明の実施例について、添付の図面を参照して詳細に説明する。

(第1の実施例)

図1は、本発明の第1の実施例における更新通知装置10の構成を示している。この更新装置10は、インターネット2及び／又は携帯電話網3を介して、複数のWebサーバ1a乃至1cと、複数の利用者端末4a乃至4cとに接続されている。Webサーバ1a乃至1cの各々は、インターネット2上にWebサイトを構成し、そのサイトアドレスの下に少なくとも1つのWebページを、インターネット2に接続された利用者端末4a乃至4cを含む多数のコンピュータ端末（図示せず）に公開するサーバコンピュータである。Webページは、一般にホームページとも称されるインターネット上に公開される情報文書であり、主にHTML(:Hyper Text Markup Language)により記述された文書である。Webページは、通常、Webサーバ1a乃至1cの各々を運用する運用者、あるいは各WebサイトやWebページの管理者により定期的又は不定期に更新される。

利用者端末4a乃至4cは、更新通知装置10により更新通知の提供を予定する利用者が使用するコンピュータあるいは携帯電話等の端末である。そして、上記更新通知は、通常は、破線で図示されるようなメールマガジンM2あるいは電子メールM1の形態により配信される。

尚、本実施例の説明では、3つのWebサーバ1a乃至1cと3つの利用者端

末 4a 乃至 4c が構成として示されているが、本発明による更新通知装置 10 の構成は、これらの数に限定されず、多数の Web サーバ及び多数の利用者端末を収容し得る。

更新通知装置 10 は、通常のサーバコンピュータを利用することが可能であり、設定部 12、更新検知部 13、通知部 15、記事生成部 14、メール作成部 17、およびデータベース（以下 DB とも称する）11 とを含む。

設定部 12 は、利用者端末 4a 乃至 4c の各々との通信により、更新されたか否かを検知する対象となる Web サーバ 1a 乃至 1c のネットワークアドレスである URL が入力されると、これを DB 11 に記録保持する。

更新検知部 13 は、DB 11 に記憶保持された URL に対応する Web サーバ 1a 乃至 1c の配信する配信する Web ページを、例えば所定の周期にてアクセスし、前回アクセス時との差異を検出する。この検出を行うために、更新検知部 13 は、差異を検出する手段として、前回アクセス時の Web ページを保持し、これと前回アクセス時の Web ページの内容とを比較し、差異がある場合にその差異部分を更新部分として抽出する機能を持っている。

記事再生部 14 は、更新検知部 13 で抽出された更新部分の内容に従って、通知文を構成する記事タイトル及び記事本文を生成する。記事タイトル及び記事本文は、例えば、1 つの Web サイトの Web ページの更新を、1 件単位として生成される。

メール作成部 17 は、記事生成部 13 により生成された 1 件乃至複数件の記事をまとめて 1 つの通知文を作成する。

通知文は、基本的には、通知先とする利用者端末 4a 乃至 4c の各々について

作成され、利用者端末 4a 乃至 4c の各々が設定部 12 を介して指定した U R L に
対応する W e b サーバ（W e b サーバ 1a 乃至 1c の何れか 1 つ又は複数）の更
新について通知する内容を含む。

あるいは、通知文を 1 つの利用者端末（例えば、利用者端末 4a）専用に作
成するのではなく、複数の利用者端末（例えば、利用者端末 4a 乃至 4c）を通
知先として想定するメールマガジンの形態であってもよい。

通知部 15 は、メール作成部 17 により作成された 1 つ又は複数の通知文を、
対応する利用者端末 4a 乃至 4c の各々にインターネット 2 又は携帯電話網 3 を
介して送信する。

データベース 11 は、更新されたか否かを検知する対象の W e b サイトある
いは W e b ページの U R L と、利用者毎のメールアドレスを格納する。尚、格
納される W e b サイトの U R L としては、W e b サイトの所謂ホームページア
ドレスのみならず、その配下に連なる特定の W e b ページを設定しても良い。

図 2 には、図 1 に示した更新通知装置 10 における更新通知の、処理手順を
示している。

先ず、更新通知装置 10 は、更新検知部 13 によって、D B 11 に設定されて
いる複数の U R L を読み込み、 $i = 1 \sim n$ (n は設定される U R L の個数を与
える正数) とする範囲でループ処理を始める（ステップ S1）。

次いで、1 つの U R L へのアクセスを行い、W e b ページデータ（以後、
今回データと称する）を新たに取得する（ステップ S2）。尚、W e b ページが
フレームを保持している場合は、フレーム内の各ページデータを保持する。

次いで、この今回データを前回アクセス時の W e b ページデータ（以後、

前回データと称する) と比較し、更新済みであれば後述のステップ S4 へ分岐し、更新がなければ後述のステップ S7 へ分岐する (ステップ S3)。

次いで、取得した今回データを、次回の比較で用いるための新たな前回データとして保存する (ステップ S4)。

次いで、更新通知装置 13 は、記事生成部 14 において、前回データと今回データの差分 (即ち、差分情報片) を取得する (ステップ S5)。この差分情報片の取得は、例えば前回データ及び今回データの各行を配列要素として格納し、前回データの 1 つの配列要素を今回データ中の全ての配列要素と比較することを、前回データの全ての配列要素について繰り返すことにより実行する (図 3 参照)。

次に、更新通知装置 13 は、記事生成部 14 において、抽出した差分を用いて記事を作成する (ステップ S6)。この記事の生成においては、例えばページタイトルを記事タイトルとし、差分情報の増分を記事本文に用いることができる。また、タイトルを差分から生成してもよく、タイトルのみ又は本文のみといった構成も可能である。

次に、更新通知装置 10 は、更新検知部 13 において、最後の U R L であるか否かを判定して、未処理の U R L があればステップ S1 へ分岐して次の U R L を処理し、未処理の U R L がなければステップ S8 へ分岐する (ステップ S7)。

次に、更新通知装置 10 は、メール作成部 17 において、ステップ S6 で生成した 1 つ以上の記事を結合し、通知文本文を作成する (ステップ S8)。この処理において、通知文本文は複数の記事を単に結合するだけでもかまわないが、

罫線に相当する記号や文字列を挿入し、見やすく構成するのが望ましい。

次に、更新通知装置 10 は、通知部 15において、生成した通知文を電子メールとして、当該 U R L を設定した利用者端末、即ち利用者端末 4a 乃至 4c の何れかに送信する（ステップ S9）。この通知文は、電子メールによる送信に代えて、印刷、F A X、記憶媒体への保存、あるいは W e b への掲載により閲覧可能な形式にて出力しても良い。この場合、D B 11 に格納されている利用者毎のメールアドレスは、出力先に応じたデータ（例えば、F A X の場合は F A X 番号）で置き換えると良い。

そして、複数の利用者がいる場合は、ステップ S1 乃至 S9 の処理を、利用者の人数分繰り返す。

図 3 は、図 2 に示された処理手順によって、差分情報を取得する方法の具体的に説明したものである。

この図 3 では、W e b ページ W 1 及び W e b ページ W 1' の各々が H T M L のタグにより表記されるテキスト形式にて示されている。W e b ページ W 1 は、更新により W e b ページ W 1' に更新されたとする。この場合、図 2 にしめしたステップ S5 の差分の取得の実行により、差分情報 $\Delta W 1$ が取得される。この図 3 の例では、差分情報 $\Delta W 1$ は、「A 社と連携（改行）文 9 ~~~~~ ~~~~~、~~~~~。（改行）文 1 0 ~~~~~。文 1 1 ~~~~~。（改行）文 1 2 ~~~、~~~~~ ~~~~~。」の如く新たに内容が追加されている。

図 4 は、通知文の生成例を示している。メールマガジン M 2 は、更新のあった W e b サイトのアドレスと当該更新の内容を通知するメールマガジンの例

である。

以上説明したように、第1の実施例では、指定されたURLに対してアクセスを行い、前回アクセス時との差分を用いて更新の通知文を生成することができる。これにより、通知文を受け取る側では、逐一更新先のURLを参照しなくとも、変更のあった記事や概要を知ることが可能となる。この結果、より詳しい情報を得るときのみ当該URLにアクセスするだけで良い。また、通知文をメールマガジンとして発行する運営者側では、Webサイトの更新にあわせたメールマガジンを作成する上で、コスト低減と時機を得た情報の発行が可能となる。

(第2の実施例)

図5は、本発明の第2の実施例の、更新通知装置10の構成を示している。この第2の実施例において、更新通知装置10に接続されている複数のWebサーバ1a乃至1c、複数の利用者端末4a乃至4c、インターネット2、及び携帯電話網3は、第1の実施例と同一の構成なので、構成についての説明は省略する。

更新通知装置10は、第1の実施例と同様に通常のサーバコンピュータを利用することができる、設定部12、更新検知部13、通知部15、記事生成部14、メール作成部17、およびDB11とを含む。本実施例の更新通知装置10は、更に、WebページのテンプレートをDB11から検索し、更新検知部13又は記事生成部14に供給するテンプレート検索部33と、メール作成部17で作成されたメールの通知文にサブジェクト、即ちメールの表題を生成するサブジェ

クト生成部 36 とを含む。

記事生成部 14 は、第 1 の実施例と同様に記事を生成するとともに、当該記事を記事テンプレートに従って生成する。メール作成部 17 は、第 1 の実施例と同様にメールを作成すると共に、サブジェクト生成部 36 から供給されるサブジェクトの内容に従って、そのメールに表題を付与する。

D B11 は、更新されたか否かを検知する対象のWeb サイトあるいはWeb ページのURL と、利用者毎のメールアドレスを格納すると共に、更に、当該Web サイトあるいはWeb ページのテンプレートを格納する。

各Web ページのテンプレートは、更新通知装置 10 の管理者が予め設定しておいてもよく、あるいは、設定部 12 により対応するURL と共に予め設定されていても良い。この場合、D B11 に格納されている利用者毎のメールアドレスは、出力先に応じたデータで置き換えると良い。そして、複数の利用者が存在する場合には、ステップ S51 乃至 S62 を繰り返す。

図 6 は、図 5 に示した更新通知装置 10 における更新通知の、処理手順を示している。更新通知装置 10 は、更新検知部 13 において、D B11 から複数の URL を読み込み、 $i = 1 \sim n$ (n は設定される URL の個数を与える正数) とする範囲でループ処理を始める (ステップ S51)。

次いで、各URL へのアクセスを行い、Web ページデータ、即ち今回データを新たに取得する (ステップ S52)。

そして、この今回データを前回データと比較し、更新済みであれば後述のステップ S54 へ分岐し、更新がなければ後述のステップ S58 へ分岐する (ステップ S53)。

更新検知部 13 は、取得した今回データを、次回の比較で用いるための新たな前回データとして保存する（ステップ S54）。

更新通知装置 10 は、記事生成部 14 において、テンプレート検索部 33 を用いて、当該 U R L に対応する W e b テンプレートを D B 11 より取得する（ステップ S55）。

次いで、当該 W e b テンプレートと更新された W e b ページデータとのマッチング及び／又は差分取得を行う（ステップ S56）。

記事生成部 14 は、このマッチングにより抽出された内容を用いて記事を作成する（ステップ S57）。尚、この第 2 の実施例においては、更新されたページ全体をテンプレートとマッチングしている形態が示されたが、テンプレートとマッチングを行い、一致した箇所のデータの差分を取る形態でも良い。

次に、更新通知装置 10 は、最後の U R L であるか否かを判定して、未処理の U R L であればステップ S51 へ分岐して次の U R L を処理し、未処理の U R L がなければステップ S59 へ分岐する（ステップ S58）。

次いで、メール作成部 17 において、ステップ S57 で生成した 1 つ以上の記事を通知文テンプレートに従って結合し、通知文の本文を作成する（ステップ S59）。ここで、通知文の本文のテンプレートは、複数の U R L に対応する更新内容を含み得る（図 7 のテンプレート T 2 を参照）。尚、テンプレート内の更新されなかった U R L の箇所は切り詰めるのが望ましい。

次に、更新通知装置 10 は、生成した通知文の発行番号を加算する（ステップ S60）。加算された発行番号は、D B 11 に記録保持される。

そして、サブジェクト生成部 36 において、表題即ちサブジェクトを生成し、

通知文本文にテンプレートに従って上記発行番号を付与する（ステップS61）。

次に、更新通知装置10は、通知部15において、生成した通知文を電子メールとして、当該URLを設定した利用者端末、即ち利用者端末4a乃至4cの何れかに送信する（ステップS62）。この通知文は、電子メールによる送信に代えて、印刷、FAX、記憶媒体への保存、あるいはWebへの掲載により閲覧可能な形式にて出力しても良い。

図7は、図5に示される構成におけるテンプレート及び通知文の例を示している。テンプレートT2は、通知文本文のテンプレートの例である。テンプレートT1は、記事のテンプレートの例である。ここで、テンプレートの機能として、任意の文字列の枠を規定する部分が、図の例のごとく「(.+?)」として記述される（図中のT1-1参照）。通知文M4は、通知文の生成例である。

以上説明したように、第2の実施例では、指定されたURLに対してアクセスを行い、当該URLに対応して設定されているWebページのテンプレートとのマッチングにより通知文を生成することができる。前述の第1の実施例では不特定多数のURLを対象としているが、この第2の実施例では、URLを限定することで、これに対応した適当なテンプレートを用いることが可能となる。この結果、このテンプレートを利用することで、可読性の高い通知文を生成することができる。

（第3の実施例）

図8は、本発明の第3の実施例の、更新通知装置10の構成を示している。

この第3の実施例において、更新通知装置10に接続されている複数のWebサーバ1a乃至1c、複数の利用者端末4a乃至4c、インターネット2、及び携帯電話網3は、第1の実施例と同一の構成なので、構成についての説明は省略する。

更新通知装置10は、第1の実施例と同様に通常のサーバコンピュータを利用する事が可能であり、設定部12、更新検知部13、通知部15、記事生成部14、メール作成部17、およびDB11とを含む。本実施例の更新通知装置10は、更に、差分情報が所定の閾値（例えば、100文字又は10行の如き値の閾値）以上の場合に差分情報より主要文を抽出する主要文抽出部24、DB11に設定されたキーワードを含まない記事を除去するフィルタ部25、利用者が注目している記事を集計する注目度集計部27、通知文のメールヘッダを生成するヘッダ生成部28、及び利用者毎の通知文発行番号をインクリメントする発行番号加算部29とを含む。

DB11は、更新されたか否かを検知する対象のWebサイトあるいはWebページのURLを格納すると共に、更に、各利用者毎に対応してキーワードを少なくとも1つ以上と、利用者毎のメールアドレスと、通知文（メールマガジン）発行番号とを含む。

図9は、図8に示した更新通知装置10における更新通知の、処理手順を示している。先ず、更新通知装置10は、更新検知部13において、DB11から複数のURLを読み込み、 $i = 1 \sim n$ （nは設定されるURLの個数を与える正数）とする範囲でループ処理を始める（ステップS21）。

次いで、各URLへのアクセスを行い、Webページデータ、即ち今回デ

ータを新たに取得する（ステップ S22）。

そして、この今回データを前回データと比較し、更新済みであれば後述のステップ S24 へ分岐し、更新がなければ後述のステップ S34 へ分岐する（ステップ S23）。

更新検知部 13 は、取得した今回データを、次回の比較で用いるための新たな前回データとして保存する（ステップ S24）。

更新通知装置 10 は、記事生成部 14 において、前回データと今回データの差分を取得する（ステップ S25）。

この差分値が上記閾値以上であればステップ S27 へ分岐し、上記閾値未満であればステップ S28 へ分岐する（ステップ S26）。

次いで、主要文抽出部 24 において、差分データに対して主要文抽出処理を行い、上記閾値未満のサイズのデータを得る（ステップ S27）。

この抽出された主要文を用いて、記事生成部 14 が記事を生成する（ステップ S28）。この記事生成は、例えばページタイトルを記事タイトルとし、差分情報の増分を記事本文に用いることができる。また、タイトルを差分データから生成してもよく、タイトルのみ、あるいは本文のみといった構成でも良い。

次に、更新通知装置 10 は、注目度集計部 27 において、当該 U R L の更新通知を希望する利用者の通知先メールアドレス数でループ処理、即ち、D B 11 から複数のメールアドレスを読み込み、 $k = 1 \sim m$ (m はメールアドレスの個数を与える正数) とする範囲でループ処理を開始する（ステップ S29）。

次いで、更新通知を希望する利用者がキーワードを設定している場合に、記事にキーワードが含まれていない場合は後述のステップ S33 へ分岐し、含ま

れている場合は後述のステップ S31 へ分岐する（ステップ S30）。

そして、通知先ごとに記事を蓄積する（ステップ S31）。

その際に、注目度集計部 27において、記事配布ランクを 1 つアップする（ステップ S32）。

最後の通知先か否かを判定して、当該更新通知を希望する利用者がループの最後であれば後述のステップ S34 へ分岐し、そうでない場合はステップ S29 へ分岐し、次の更新通知を希望する利用者について処理を行う（ステップ S33）。

最後の U R L か否かを判定して、未処理の U R L があればステップ S21 へ分岐して次の U R L を処理し、未処理の U R L がなければ後述のステップ S35 へ分岐する（ステップ S34）。

次に、更新通知装置 10 は、更新通知を希望する利用者ごとの処理、即ち、D B 11 から複数のメールアドレスを読み込み、 $k = 1 \sim m$ (m はメールアドレスの個数を与える正数) とする範囲でループ処理を開始する（ステップ S35）。

注目度集計部 27においてランク付けした記事の、例えば上位ランク 10 件を抽出する（ステップ S36）。

次いで、メール作成部 17 において上記上位 10 件を先頭に、残りの記事を結合し、通知文本文を作成する（ステップ S37）。

これらのステップ S36 及びステップ S37 では、すべての記事をソートしてランク順に結合しても構わない。また、通知文本文生成に際しては、単に複数の記事を結合するだけでも構わないが、罫線に相当する記号や文字列を挿入し、利用者が見易い構成とすることが望ましい。

次に、更新通知装置 10 は、発行番号加算部 29において、更新通知を希望する利用者ごとに、発行番号を加算する（ステップ S38）。

次いで、ヘッダ生成部 28において、生成した通知文本文に上記発行番号を付与したサブジェクト及び通知先のメールアドレスを付与して、ヘッダを生成する（ステップ S39）。このサブジェクトは、固定文字列でもよいし、タイムスタンプや、記事中から任意あるいはランダムに抽出した記事タイトルや本文を用いてもよい。また、発行番号はサブジェクト中にある必要はなく、本文中に含まれる構成としてもよい。

次に、更新通知装置 10 は、通知部 15において、通知文のヘッダに従い、更新通知を希望する利用者のメールアドレスへ、生成した通知文を送付する（ステップ S40）。

そして、処理中の更新通知を希望する利用者が最後の利用者であるか否か判定し、最後の利用者であれば処理を終了し、最後でなければステップ S35へ分岐して次の利用者について処理を継続する（ステップ S41）。

尚、ステップ S40 の電子メールによる通知文の送信に代えて、印刷、FAX、記憶媒体への保存、あるいはWebへの掲載により閲覧可能な形式にて出力しても良い。この場合、DB11 に格納されている利用者毎のメールアドレスは、出力先に応じたデータ（例えば、FAXの場合にはFAX番号）で置き換えると良い。

図10は、メールマガジンM3の生成例を示している。ここで、メールM3は、複数のブロックM31、M32及びM33を含んでいる。ブロックM31には、加算されていく発行番号が示され、ブロックM32にはランキングが

示され、そしてブロックM 3 3には、その他の更新記事が示される。

以上説明したように、第3の実施例では、利用者毎に更新通知を希望するURLとキーワードを指定しておくことで、各利用者は、当該利用者の希望に従った更新内容が記載された通知文を受け取ることが可能となる。そして、このメールマガジンでは、更新内容が主要文に変換されて利用者の容易な理解を可能となる。更に、ランキング情報として他の利用者がどのような記事に注目しているのかの情報も掲載可能であり、メールマガジンとしての完成度を高めることが可能である。

本発明による更新通知装置は、インターネットやイントラネットに接続されたシステムへの組み込みや適用が可能であり、更新通知のサービスサイトを運営することも可能である。そして、本発明の各構成要素は、ソフトウェアあるいはハードウェアの何れの構成としても実現することが可能である。

以上のように本発明による更新通知装置及び方法によれば、Webページの更新の有無の出力のみならず、その更新内容又はその主要文が通知文に付加されて利用者に出力される。従って、登録されたWebページに、利用者にとって有意な内容の更新があったか否かを容易に判断し得る。